

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently Amended) A processor-implemented ~~non-iterative~~ method of clustering a set of records, each of the records having attribute values for a set of attributes, the method comprising:

performing a first pass across the set of records, the first pass including:

for each attribute of the set of attributes, determining a characteristic value for said each attribute, the characteristic value being one of a mean value and a median value of the attribute values of said attribute across the records;

performing a second pass across the set of records, the second pass including:

for each attribute value, determining a deviation from the characteristic value of said each attribute; and

for each record, sorting the set of attributes based on the absolute values of the deviations of the attribute values from the characteristic value of said each attribute from largest to smallest, to provide a key, the key comprising an ordered list of the set of attributes and the deviations from the characteristic value of said each attribute, the key thereby showing which attributes in the set of attributes deviate the most from the characteristic value; and

combining the set of records based on the key into a clustering result that includes a plurality of clusters;

~~wherein the key comprises an ordered list of the set of attributes and the deviations from the characteristic value of said each attribute; and~~

refining the clustering result by:

identifying a cluster having a smallest number of records;

for each record of the identified cluster, searching another cluster having records with best matching keys, based on a user-provided criterion; and

distributing the cluster with the smallest number of records to the other cluster having records with best matching keys, to reduce the total number of clusters; and

repeating the identifying, searching and distributing steps until a desired number of clusters has been reached.

2. Canceled.
3. Canceled.
4. (Original) The method of claim 1, wherein determining the deviation comprises calculating a difference between said each attribute value and the characteristic value of said each attribute.
5. (Original) The method of claim 1, wherein determining the deviation comprises calculating a difference between said each attribute value and the characteristic value of the corresponding attribute, and dividing the difference by the characteristic value of said each attribute.
6. Canceled
7. (Original) The method of claim 1, wherein a first record of the set of records contains a first key and a second record of the set of records contains a second key; and
further comprising placing the first key and the second key into a single cluster if the first key and the second key have identical subsequences of a first length.
8. (Original) The method of claim 1, wherein a first record of the set of records contains a first key and a second record of the set of records contains a second key; and
further comprising placing the first key and the second key into a single cluster if the first key and the second key have identical subsequences of absolute values of the deviations.
9. (Original) The method of claim 1, wherein a first record of the set of records contains a first key that has a first sub-sequence, and a second record has a second sub-sequence contains a second key; and
further comprising placing the first key and the second key into a single cluster if the first and second sub-sequences comprise the same set of attributes.

10. (Original) The method of claim 9, wherein the first and second subsequences comprise the same set of attributes irrespective of a sign of the deviations of the attribute values.
11. Canceled.
12. (Previously Presented) The method of claim 9, further comprising reducing a length of the first sub-sequence and a length of the second sub-sequence in order to find a best match.
13. (Original) The method of claim 12, further comprising using a distance measure to find another cluster for a record of the identified cluster.
14. (Previously presented) The method of claim 13, wherein the distance measure comprises a Euclidean distance.
- 15-26. Canceled